The Effects of Illumination Direction on the Perception of 3D Shape from Shading

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation *with honors research distinction* in Psychology in the undergraduate colleges of The Ohio State University

by

Makaela S. Nartker

The Ohio State University
April 2016

Project Advisor: Dr. James T. Todd, Department of Psychology

Abstract

A fundamental problem for the perception of 3D shape from shading is to achieve some level of constancy over variations in the pattern of illumination. The present research was designed to investigate how changes in the direction of illumination influence the apparent shapes of surfaces. The stimuli included images of 3D surfaces with Lambertian reflectance functions that were illuminated by a rectangular area light source. The direction of illumination was systematically manipulated. Observers judged the 3D shapes of these surfaces by marking local depth minima and maxima along three designated scan lines using a hand-held mouse. The results revealed that the local depth maxima were shifted slightly toward the direction of illumination, while the local depth minima were shifted slightly away from the direction of illumination. However, these changes were much smaller than what would be expected based on differences in the pattern of luminance among the stimulus images. These findings demonstrate that there is a substantial amount of illumination constancy in the perception of 3D shape from shading, but that it is not perfect. Several hypotheses are considered about how this constancy could potentially be achieved.

Keywords: 3-D shape, perception, shading, vision

Introduction

Human observers have the remarkable ability to perceive 3-dimensional shapes from 2-dimensional patterns of image shading. The human visual system integrates a multiplicity of cues in order to provide us with information about depth. Some of these cues include shading, texture, stereo, and motion. However, shading can create compelling impressions of 3D shape when it is the only cue available. Artists have known for a long time that shading can elicit perceptions of 3D form (see Figure 1 for examples). Leonardo da Vinci helped perfect a shading technique known as chiaroscuro (from Italian *chiaro,* “light,” and *scuro,* “dark”), which has been used for decades as a primary means of depicting depth in art. Da Vinci did experiments with light and shading, and he wrote in his famous *Notebooks* that shading was of “supreme importance” for the realistic depiction of objects in space (Todd & Mingolla, 1983, p. 583).

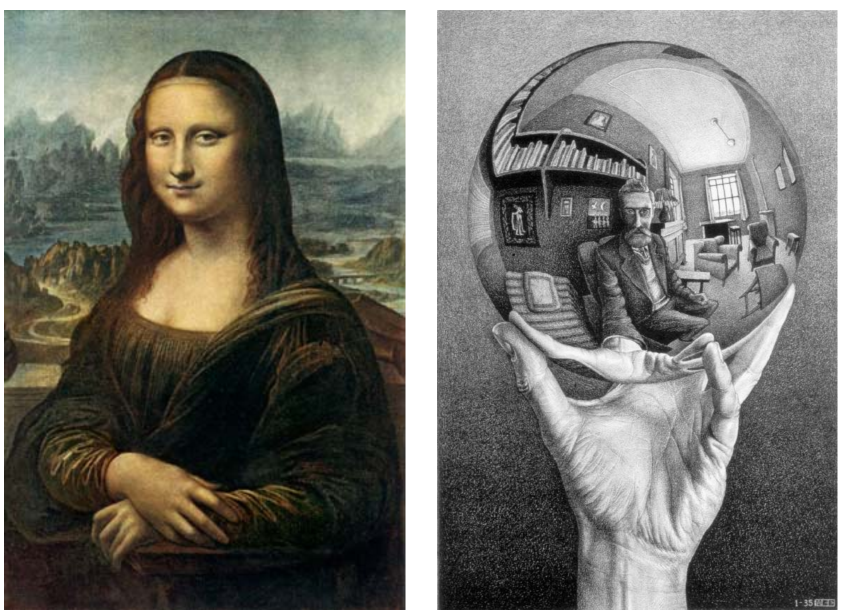

Figure 1. Shading can be a strong cue for the perception of 3D shape, even in 2-dimensional images. *Mona Lisa* (left) by Leonardo da Vinci (1503-1506), and *Hand with Reflecting Sphere* (right) by M.C. Escher (1935).

Despite the long history of shading in art, vision scientists still seek to explain how the human visual system uses shading as a cue to inform us about depth. Many models have been proposed in the literature for how shape from shading is accomplished, but their performance has been disappointing relative to that of human observers. Interpreting patterns of shading is difficult, because the light that reflects off a surface toward an observer is determined by three factors: (1) the pattern of local surface orientation (i.e. shape), (2) the material properties of the surface (e.g. albedo), and (3) the pattern of illumination. How does the visual system untangle these three sources of information? This problem is more difficult than it may seem, as there are an infinite number of possible combinations of shapes, materials, and light sources that can exactly reproduce a single image. This is sometimes referred to as the many-to-one mapping problem. This inherent ambiguity of shaded images has led many researchers to conclude that the human visual system is performing a statistical inference. In other words, the system isn't necessarily determining the real shape, but rather the *most likely* shape. Because of this, shape from shading models have to make simplifying assumptions or add additional viewpoints in order to output a single solution for the shape of an object in a shaded image.

In most studies, the light that reflects from a local surface region towards the position of observation (i.e. luminance) varies systematically with surface orientation. This is typically descried using the bi-directional reflectance distribution function (BRDF; Nicodemus, Richmond, Hsia, Ginsberg, & Limperis, 1997; Todd, Egan & Phillips, 2014). For a given homogenous illumination and viewing direction, the BRDF describes a specific mapping between local surface orientation and luminance. Under homogenous illumination, all points with the same local 3D orientation must always have the same luminance (Todd, Egan, & Phillips, 2014). However, without the assumption of homogenous illumination, it is possible for local regions with different 3D orientations to have the same luminance (hence the many-to-one mapping problem). Pont and Koenderink (2007) proposed four theoretic BRDFs that repserent generic types of surface materials found in the natural environment. The most common BRDF used in psychophysics experiments is diffuse or Lambertian reflectance function on matte surfaces. Specular surfaces complicate the analysis of image shading because they reflect light primarily in one direction, which results in reflections of other objects in the surrounding environment (M.C. Escher's chrome sphere in Figure 1 illustrates this well). Surfaces with Lambertian reflectance functions diffuse light equally in all directions. Because of this, the intensity of the corresponding picture element can be simulated using Lambert's law: $I_P = I_{LS}$(L-N), where I_P is the intensity of the light source; s is the shade or albedo of the surface ranging from 0 (black) to 1 (white); L is a unit vector in the direction of the light source; N is a surface normal (i.e. a unit vector that is perpendicular to the surface). Lambert's law says that the luminous intensity observed from a reflecting surface is directly proportional to the cosine of the angle between the direction of incident light and the surface normal.

Previous research has shown that direction of illumination has no detectable effect when stimuli include additional cues, such as texture or binocular stereopsis (Todd, Norman, Koenderink, & Kappers, 1997). However, for more impoverished stimuli, such as the ones created for this experiment, illumination changes have been shown to have a dramatic effect on perceived structure (Mingolla & Todd, 1986). Several other studies have suggested that changes in illumination direction cause systematic distortions in perceived 3D structure (Caniard & Fleming, 2007; Christou & Koenderink, 1997). However, only one other study used stimuli with realistic illumination. They used a single point light source, which almost always creates impoverished viewing conditions (Egan, 2014). The current experiment is unique in two ways:

first, we use software that realistically approximates the way that light interacts with surfaces in the natural world; second, we systematically manipulate the direction of illumination. The goal of the current experiment is to determine the effects of illumination direction on the perception of 3D shape, and discuss the implications for current shape from shading models.

Materials and Methods

Participants

Six observers participated in the experiment, including two authors and four others who were naïve about the issues being investigated. All observers had normal or corrected-to-normal visual acuity. Participants wore an eye patch over one eye to eliminate conflicting flatness cues from binocular vision.

Apparatus

The experiment was conducted using a Dell Precision 1650 PC with an NVIDIA Quadro 600 graphics card. The stimulus images were presented on a 10-bit, 28-in. gamma-corrected LCD with a spatial resolution of 2560 · 1440 pixels. The images were displayed within a 32.5-cm · 32.5-cm region (1024 · 1024 pixels) of the display screen, which subtended 18.58 · 18.58 of visual angle when viewed at a distance of 100 cm.

Stimuli

The stimuli consisted of 21 images of three smoothly curved surfaces with Lambertian reflectance functions. The images were generated in 3D Studio Max; a camera, which captures each image frame, was fixed in front of the image at the positioned labeled 0° (see below). A rectangular area light source was rotated from 60° to the left to 60° to the right in 20° increments, making a total of seven illumination directions for each surface. The illumination conditions will be referred to as -60° (to the left), -40°, -20°, 0°, 20°, 40°, and 60° (to the right). Figure 2 illustrates the stimuli presented in the experiment.

The stimuli were rendered using Maxwell Render, an unbiased rendering software that realistically approximates the way that light interacts with surfaces. Systematic manipulation of the light source and the use of an unbiased renderer were crucial to the experimental design. Previous experiments investigating the effects of illumination have used stimuli created with biased renderers, which allow an individual to exert manual control over certain settings; for example, limiting the number of bounces that the light makes in the interest of efficiency. In contrast, an unbiased renderer such as Maxwell continues to run, converging on the "correct" solution. Because of the computational power required, the images used in the present experiment took more than 10 minutes each to render on a 128 node cluster. A better understanding of how light interacts with surfaces could be applied to speeding up rendering algorithms and creating more realistic computer graphics (see *discussion* for more practical applications).

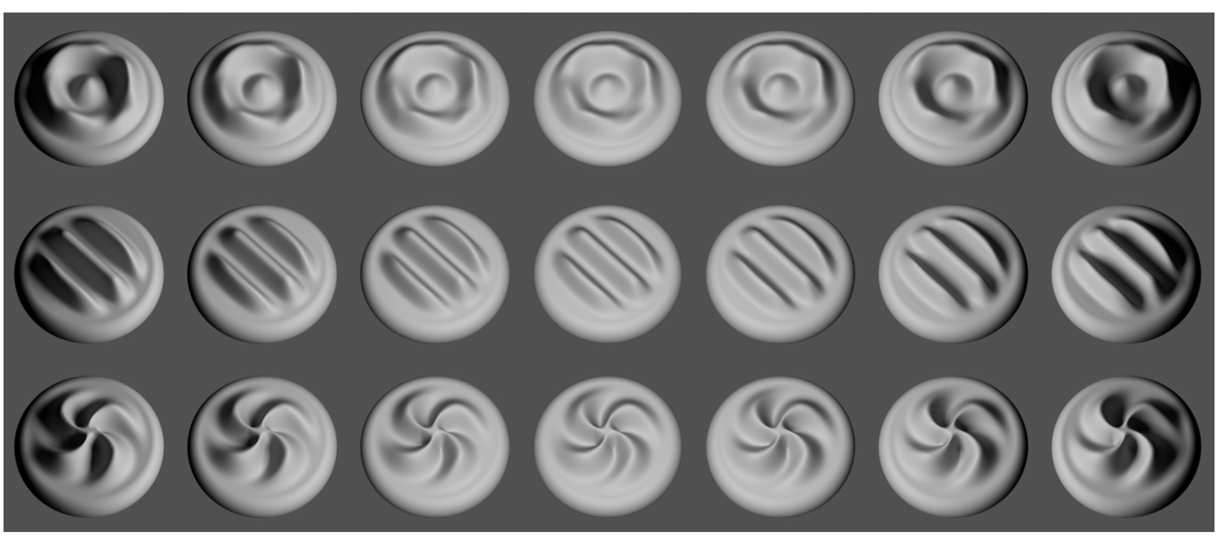

Figure 2. The stimuli consisted of 21 images of smoothly-curved surfaces. The images were generated using 3D Studio Max and Maxwell Render. All of the stimuli were presented once during an experimental session.

Procedure

The task used for the present study required observers to make local shape judgments for each of the stimuli. Participants were asked to identify local depth extrema (minima and maxima) along three designated horizontal scan lines for each of the stimuli. On each trial, an image was presented on the screen and participants marked local near and far points using a hand-held mouse. The mouse was restricted so that it could only be moved horizontally along a given scan line; a left click with the mouse dropped a blue dot, or a far point (local depth maxima), and a right click dropped a yellow dot, or a near point (local depth minima). The scan lines were always located at 360 pixels, 600 pixels, and 840 pixels on the images, which were 1200 x 1200 pixels (see Figure 3). Once an observer was satisfied with their responses for a given line, pressing the CONTROL button on the keyboard terminated the trial and a new display was presented.

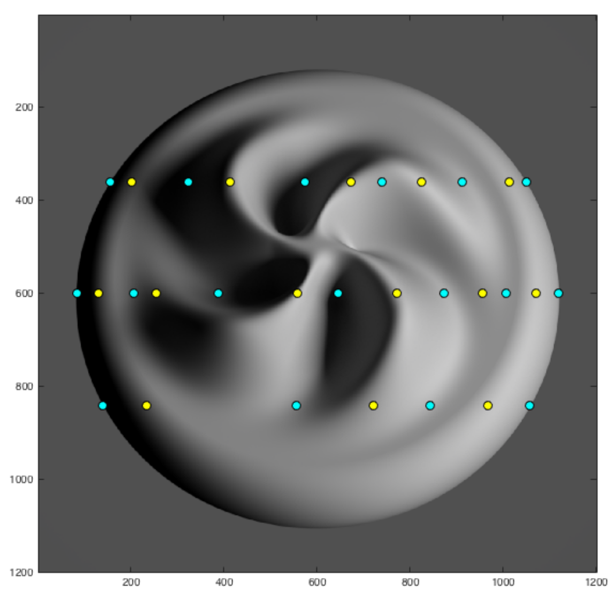

Figure 3. **Task.** An example of responses from a single participant for a single stimulus. Note: the numbers along the axes have been added to illustrate the image dimensions (in pixels) and locations of the three scan lines, but they were not present during the actual experiment. Observers were instructed to mark local depth maxima (near points) with a yellow dot, and local depth minima (far points) with a blue dot.

Results

Performance on the task was measured by comparing the location of observers' responses to the corresponding ground truth extrema. First, a simple signal detection analysis was performed and judged points were separated into either hits, misses, or false alarms. There were between three and five legitimate minima and maxima along each scan line, and every participant succeeding in marking all (100%) of the legitimate extrema on each surface. The category of false alarm denotes marked points that are neither minima nor maxima. Due to the topological constraints on the distribution of near and far points, false alarms typically come in pairs of one near point and one far point (see Figure 4). The rate of false alarms was high, with some pairs of false alarms marked by every participant for every stimulus.

Further conclusions can be drawn from interpreting the rates and locations of false alarms. The edges of each stimulus appear to be raised, and this region consistently elicited false alarm responses from observers. The stimuli were slanted back slightly because fronto-parallel images can be somewhat unstable – this tilt may be responsible for producing the pattern of shading that caused the impression of a raised edge. Alternatively, this phenomenon could be caused by shadows being cast from the surface features themselves (i.e. the "hills" and "valleys"). Whatever the cause, the high rate of false alarms indicates that while human perception is remarkably stable, it is not always entirely accurate. Determining when and how the human visual system errs will provide valuable insight into the mechanisms it may be employing to solve shape from shading.

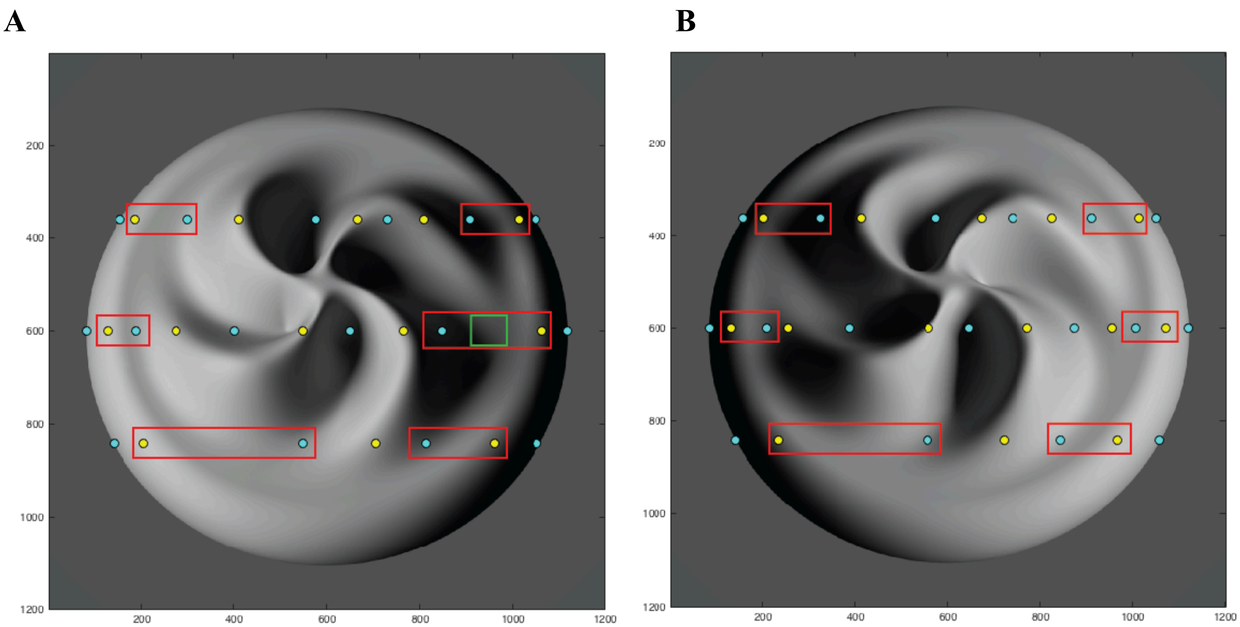

Figure 4. **False alarms.** Red boxes indicate false alarm pairs. Green boxes indicate missed extrema. On the right side of the second scan line in Image B, notice two additional points marked by observers that are not marked in Image A. These points are hits in Image B, but one is missed in Image A. The illumination condition caused participants to miss an extrema in one condition (left), and add another false alarm in the other (right).

The standard error of the responses for each extrema location were calculated. In an attempt to visually represent the variance, error bars were added to both the images and scan line plots, but the variance was so small that the bars were not even visible. Figure 5 below represents the results for the -60° and 60° conditions of this surface. The lines on the right represent the responses for the image to the left. There are three rows for each image, one for each of the three scan lines. The luminance profile is the brightness (0 = black, 1 = white) along a given scan line. The depth profile is the actual curvature of the surface, and the peaks and valleys of this profile illustrate the actual minima and maxima. Thus, the depth profile is the same for each condition of the three surfaces. The difference in illumination direction causes dramatic changes to the luminance profiles for each scan line. However, the location of points marked remains almost unchanged. The shift is so small that there are few visible differences between the responses for the two conditions.

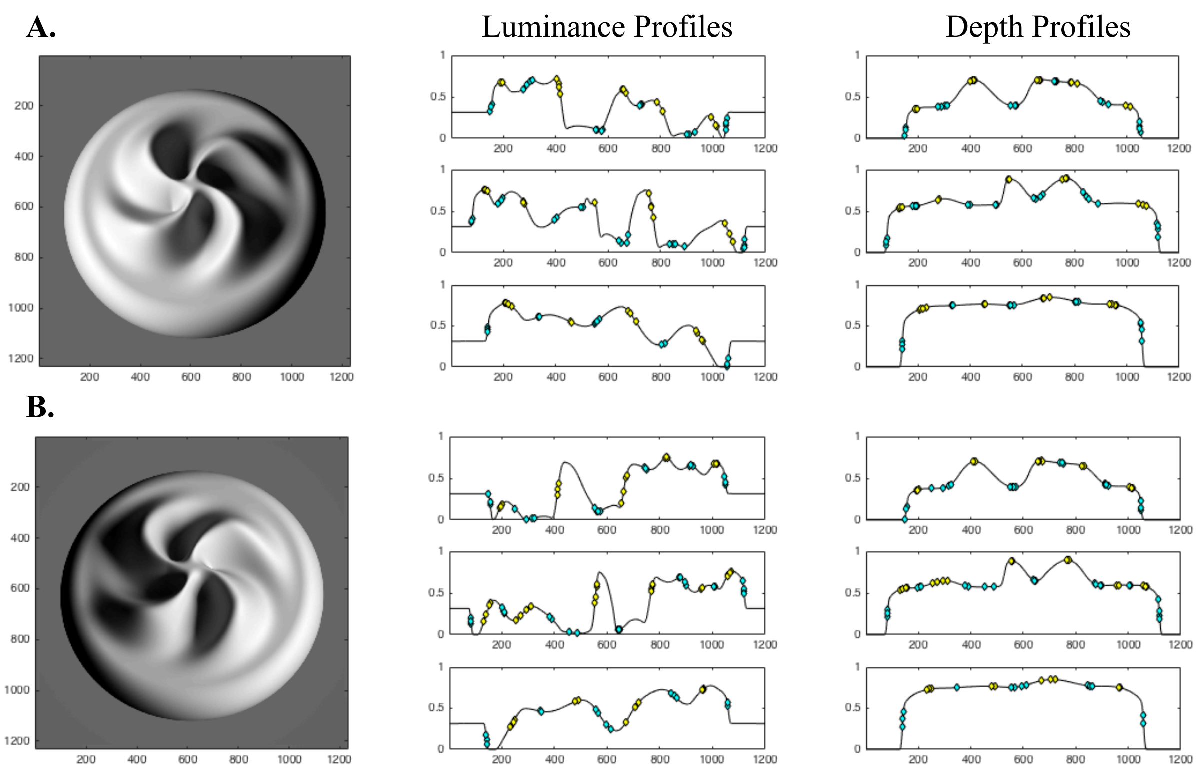

Figure 5. **Panel A** represents the responses for this stimulus in the -60° illumination condition. **Panel B** represents the responses for the same stimulus when illuminated from 60° to the right. The column labeled 'Luminance Profiles' represents the brightness along a given scan line. The column labeled 'Depth Profiles' represents the actual curvature along that line. Changing illumination direction produced large changes in the luminance profiles, but the locations of judged near points was relatively unchanged.

All of the responses for a given extrema point were averaged, and the mean response was subtracted from the ground truth (veridical) location of that extrema on the surface. A correlation between the illumination condition and the location of judged near points revealed a slight shift of judged near points toward the direction of illumination. The correlation between the average location of near points and the direction of the light source was 0.978. So while observers displayed a very high degree of shape constancy, there was a slight systematic shift of about 10 pixels (very small in terms of visual angle) toward the direction of illumination (see Figure 6).

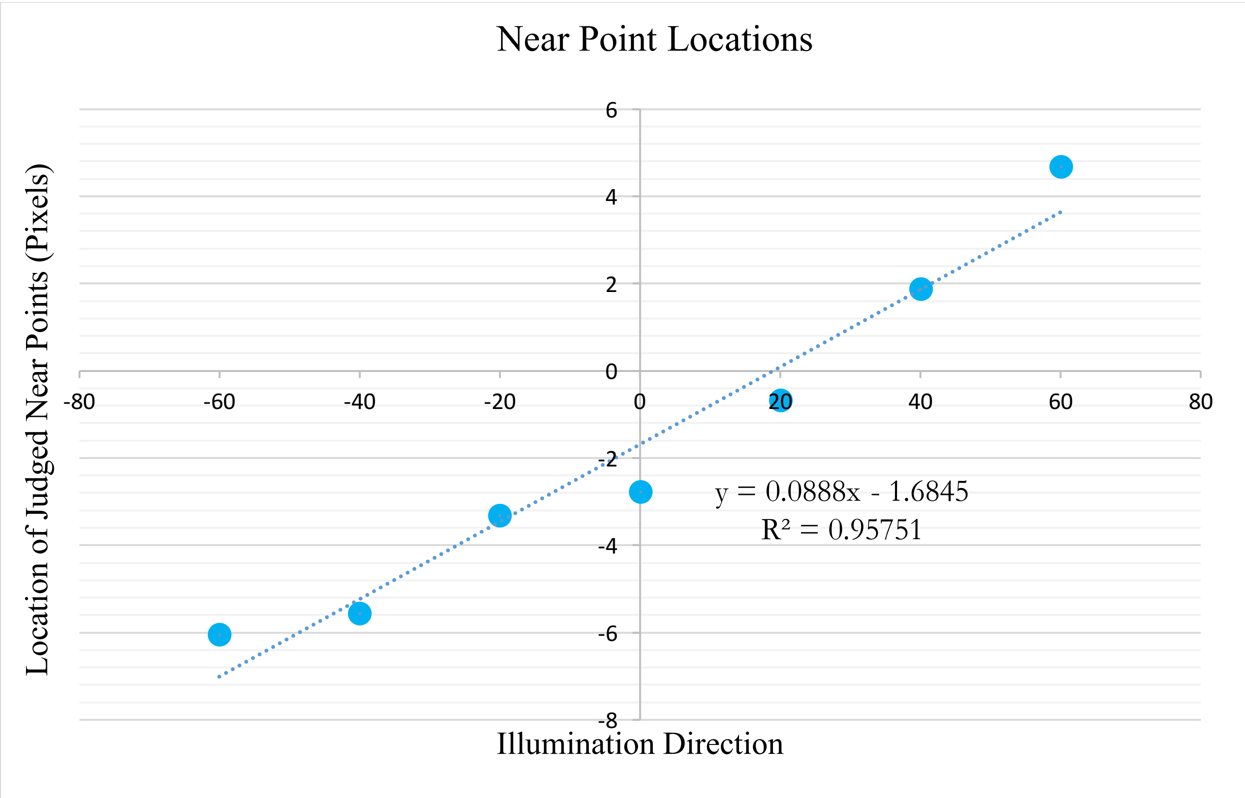

Figure 6. The graph represents the correlation between direction of illumination and average location of judged near points. As the light source moved from 60° to the right to 60° to the left of the surface, the average location of judged near points shifted by about 10 pixels.

The same analysis was performed for the average location of far points, and the correlation demonstrated that local minima (far points) shift away from the direction of illumination (see Figure 7). As the light source changed direction from left to right, far points shifted in the opposite direction, away from the light source. The correlation for far point locations was r = -0.80, smaller than that of the near points and in the opposite direction. The lower correlation coefficient may be a result of the fact that far points tend to be in locations where there were shallow shading gradients, which led to more variation. Near points on the other hand tend to be around steep gradients, which explains why there is less variance in near point responses than far point responses.

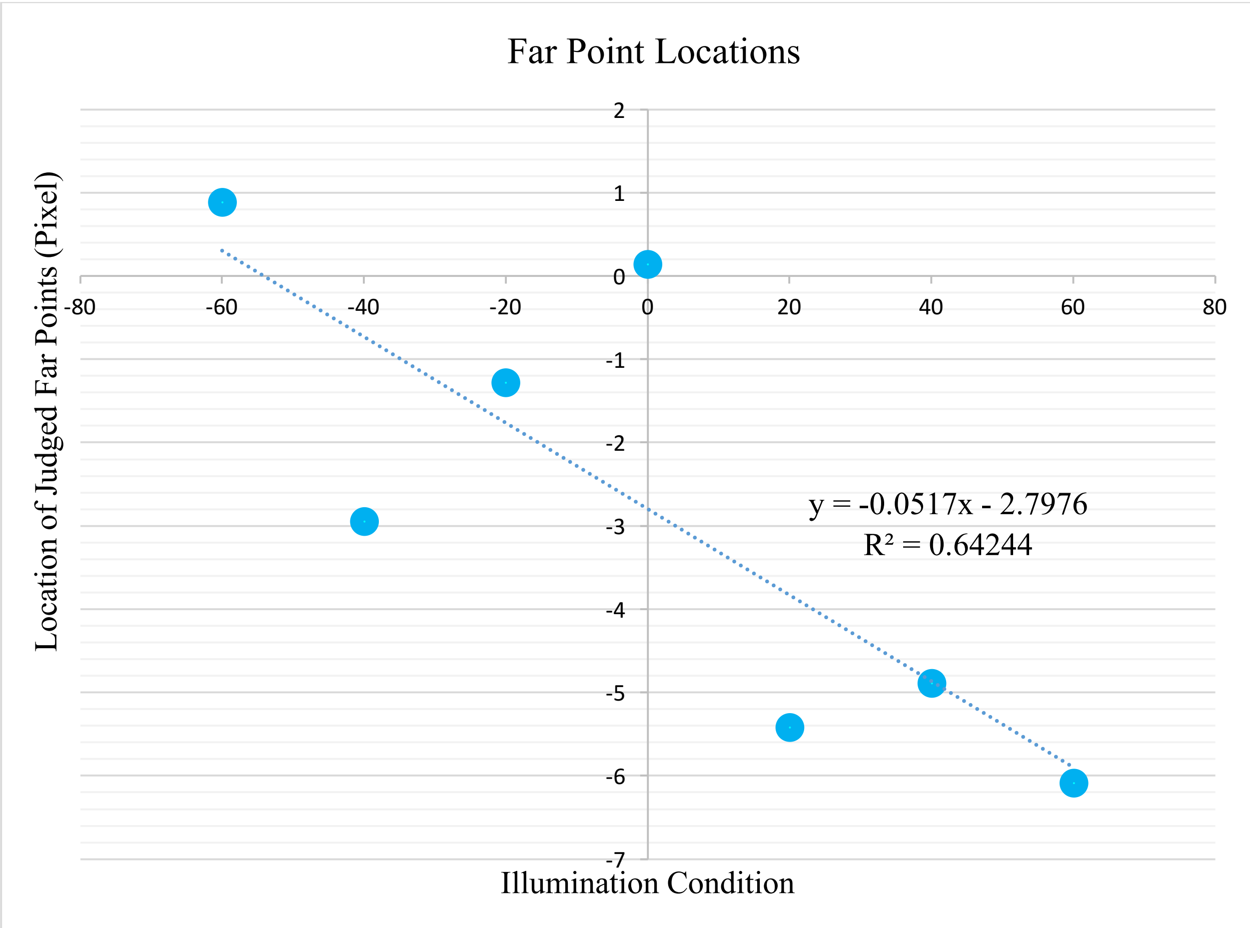

Figure 7. The graph represents the correlation between direction of illumination and average location of judged near points. As the light source moves from 60° to the right to 60° to the left of the surface, the average location of judged far points shifted by about 8 pixels.

Discussion

These results demonstrate that while there is a systematic shift caused by changing the illumination direction, these changes are much smaller than what would be expected based on differences in the pattern of luminance across the stimulus images. This suggests that consistent with previous research, there is a substantial amount of illumination constancy in the perception of 3D shape from shading, but that it is not perfect. Varying the direction of illumination causes minute, but detectable systematic shifts in apparent shape. Interestingly, varying the light source direction has different effects for different surface features. Near points tended to shift toward the direction of illumination while far points shifted in the opposite direction. It seems likely that light source direction also has a systematic effect on the prevalence and location of false alarms, but the analysis completed for this paper is not sufficient to determine the nature of these shifts.

The human visual system clearly has a remarkable mechanism for integrating information about illumination to information about surface shape and reflectance. Otherwise, human observers would have a difficult time maintaining stable interpretations of 3D shape across different patterns of illumination. However, this mechanism must be rather general in order to consistently arrive at a single interpretation. This research has important implications for shape from shading algorithms, because it suggests that the assumption of homogenous illumination may not be psychologically valid. In addition, the results underscore the importance of using realistic illumination in psychophysics experiments.

Practical Applications

There are several ways in which this line of research can be practically applied. For example, O'Hara & Barnes (2012) applied a shape from shading algorithm to images captured by the Mars Express High High Resolution Stereo Camera (HRSC). In this case, shape from shading was used to refine models created using stereo techniques. Determining how the pattern of illumination direction affects apparent shape could also be applied to facial recognition software, so that particular faces can be reliably recognized across variety of illumination conditions. Lastly, this research can be applied directly to improving computer graphics software by both speeding up rendering time and creating more realistic computer graphics.

Future Work

Having demonstrated that illumination direction causes systematic changes in perception, further research should investigate the effects of illumination direction on surfaces with non-Lambertian reflectance functions. While this study focuses primarily on the effectiveness of shading as a cue to shape, careful interpretation of the results may also give insight into which, if any, of the simplifying assumptions adopted by the computer vision community to overcome problems of image analysis are employed by the human visual system.

References

Caniard, F., & Fleming, R. (2007). Distortion in 3D shape estimation with changes in illumination. In C. Wallraven & V. Sunstedt (Eds.), Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization, 253, 99–105.

Christou, C. G., & Koenderink, J. J. (1997). Light source dependence in shape from shading. Vision Research, 37(11), 1441–1449.

Egan, E. J. (2014). Ph.D. Dissertation. The role of illumination direction on the perception of three dimensional shape from shading. *Department of Psychology*, Ohio State University.

Egan, E. J., & Todd, J. T. (2015). The effects of smooth occlusions and directions of illumination on the visual perception of 3-D shape from shading. Journal of Vision, 15, 1-11. doi:10.1167/15.2.24.

Egan, E., Todd, J. T., & Phillips, F. (2012). The role of symmetry in 3D shape discrimination across changes in viewpoint. *Journal of Vision* 2012;12(9):1048. doi:10.1167/12.9.1048.

Mingolla, E., & Todd, J. T. (1986). Perception of solid shape from shading. Biological Cybernetics, 53, 137–151.

Koenderink, J. J., van Doorn, A. J., Christou, C., & Lappin, J. S. (1996a). Shape constancy in pictorial relief. Perception, 25, 155–164.

Koenderink, J. J., van Doorn, A. J., Christou, C., & Lappin, J. S. (1996b). Perturbation study of shading in pictures. Perception, 25, 1009–1026.

Langer, S. L., & Bülthoff, H. (2000). Depth discrimination from shading under diffuse lighting. Perception, 29, 649-660. doi:10.1068/p3060

Todd, J. T., Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. (1996). Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. Journal of Experimental Psychology: Human Perception and Performance, 22(3), 695–706.

Todd, J. T., & Mingolla, E. (1983). Perception of surface curvature and direction of illumination from patterns of shading. Journal of Experimental Psychology: Human Perception and Performance, 9(4), 583–595.